four standard dideoxynucleotide terminators. The base periodicity is the mass of dN, or 310 daltons. Figure 2b shows a target second sequence resolved on the same mass spectrum shown in Figure 2a, using a primer heavier by 77 daltons. The peaks corresponding to the reaction products from the first target sequence can fall within the spectrum in Figure 2b, which can never intersect peaks from the second target sequence. This permits unambiguous resolution of both sequences each peak can be uniquely assigned to a nucleotide, a base position, and a target sequence.

Figure 3 shows four different sequences resolved in a single spectrum using a set of mass-staggered primers that are separated in mass by integer multiples of 77 daltons (77, 154, and 231 daltons).

A1

Figure 4a-d shows the general implementation of a forced mass modulation method using pair-matched nucleotides, for the analysis of sequencing reaction products as double-stranded structures. The steps in the reaction are as follows: a) a partially duplex hairpin primer with a 3′ overhang and a 5′ phosphate group is annealed and ligated to the single stranded target sequence; b) the resulting partially duplex structure is subjected to a sequencing reaction using the pair-matched nucleotide set described above along with the set of mass-matched terminators (ddM); c) products resulting from sequencing reaction b); and, d) the products c) from the sequencing reaction are exposed to a strict single strand-specific nuclease that results in the production of blunt-ended hairpin structures ready for analysis by mass spectrometry.

Figure 5a-b shows the products and molecular masses of the nuclease digestion elucidated in Figure 4d, along with a simulated mass spectrum.

Figure 6a-d shows three sequence variants (Figure 6a) that differ from each other only at a single base position sequenced by a conventional Sanger reaction. Figure 6b is a simulated mass spectrum of all reaction products shown in Figure 6a. Figure 6c is a graph representing the valid sequence permutations that can be elucidated from the mass spectrum shown in Figure 6b. Boxed values are fragment masses, solid arrows show valid sequence branches, dashed arrows represent

spurious branches. In practice valid branches are indistinguishable from spurious ones. Figure 6d is a set of sequences consistent with the graph shown in Figure 6c. Spurious sequence reconstructions are shown in lowercase letters, valid ones in uppercase letters.

Figure 7a shows the three related sequence variants from Figure 6a-d sequenced by Forced Mass Modulation using a single primer and the mass-matched nucleotide set from Figure 2a-b with the standard dideoxy terminators. The positions of the differing bases are shown by solid arrows. Reaction products are shown along with their respective molecular masses. Reaction products of variant #2 whose masses differ from those of variant #1 are marked with by (*). Reaction products of variant #3 whose masses differ from those of variant #1 are marked by (**). Figure 7b is a simulated mass spectrum of all reaction products shown in Figure 7a along with a sequence graph. The dotted regions represent the only valid mass ranges that can be assumed by the reaction products from Figure 7a. The base periodicity is 310 daltons. Figure 7c is a consensus sequence derived from the data shown in Figure 7b. Figure 7d is an expansion of the consensus sequence shown in Figure 7c. Spurious reconstructions are shown in lowercase letters, valid ones in uppercase letters. Note that there is only a single spurious reconstruction, as opposed to the eleven errant sequences reconstructed from the Sanger reaction described in Figure 6a-d.

A1

Figure 8i-iii shows the base composition density distributions for the total set of possible 7-base oligonucleotides using three different nucleotide sets. Note that for the set of naturally occurring bases, nearly every base composition has its own distinct mass value, but most of these mass values are spaced only one dalton from each other. Increasing the peak separation also markedly increases the average number of base compositions per observed mass, particularly for those masses in the center of the range.

Please replace the paragraph on pages 15, line 29, to page 16, line 13, with the following:

As used herein, the term "mass-matched nucleotides" refers to a set of nucleotide analogs wherein each analog is of identical mass to each of the other analogs. For example, analogs of dA, dG, dC and dT can form a mass-matched nucleotide set, when each analog is selected to have the same molecular weight as the others in the set. Mass-matched nucleotide sets can be identified by selecting chemically modified derivatives of natural bases or by the use of a universal base analog such as deoxyinosine or 5-nitroindole and 3-nitropyrrole (5-nitroindole and 3-intropyrrole can be in the dideoxy form) which can form base pairs with more than one of the natural bases. Others include, 3-methyl 7-propynyl isocarbostyril, 5-methyl isocarbostyril, and 3-methyl isocarbostyril. As a result, oligonucleotides that contain such bases differ in molecular weight only as a function of length thereof. Furthermore, incorporation of a single nucleotide(s) that is (are) not in the set renders such the oligonucleotide(s) readily identifiable by mass, particularly by spectrometric analysis.

Please replace the paragraph on page 19, lines 5-19, with the following:

As used herein, a "primer" refers to an oligonucleotide that is suitable for hybridizing, chain extension, amplification and sequencing. Similarly, a probe is a primer used for hybridization. The primer refers to a nucleic acid that is of low enough mass, typically about between about 5 and 200 nucleotides, generally about 70 nucleotides or less than 70, and of sufficient size to be conveniently used in the methods of amplification and methods of detection and sequencing provided herein. These primers include, but are not limited to, primers for detection and sequencing of nucleic acids, which require a sufficient number nucleotides to form a stable duplex, typically about 6-30 nucleotides, about 10-25 nucleotides and/or about 12-20 nucleotides. Thus, for purposes herein, a primer is a sequence of nucleotides of any suitable length, typically containing about 6-70 nucleotides, 12-70 nucleotides or greater than about 14 to an upper limit of about 70 nucleotides, depending upon sequence and application of the primer.

Please replace the paragraph on page 20, lines 16-19, with the following:

A4

As used herein, pattern with reference to a mass spectrum or mass spectrometric analyses, refers to a characteristic distribution and number of signals (such peaks or digital representations thereof).

Please replace the paragraph on page 30, lines 2-19, with the following:

As noted above, Forced Mass Modulation refers to methods provided herein that permit unambiguous assignment of peak positions (or masses) to mass fragments of oligonucleotides according to their base position, base identity, and target sequence from which the fragments arose. The methods use deoxynucleotide analogs, modified nucleotide terminators, mass-labeled primers, mass-staggered primers and other such nucleotides, nucleic acids and analogs thereof to provide a means for deconvoluting complex mass spectra or output from other mass determining techniques. These methods permit deconvolution of highly multiplexed nucleic acid reaction mixtures for sequencing methods and detection methods that include a step of primer extension. In practicing these methods, primers are extended using mass-matched nucleotides and chain terminators (or in some embodiments mass where it is only necessary to detect incorporation (or the absence of incorporated) mass-matched terminators and optionally mass-matched chain extending nucleotides). Because the sequence and/or molecular mass of a primer is known, and the extended nucleotides have the same molecular mass, a periodicity in molecular mass that is a function of molecular weight of the selected mass matched nucleotide(s) results.

A5

Please replace the paragraph on pages 41, line 14, to page 42, line 6, with the following:

The computer can be part of the instrument or another system component or it can be at a remote location. A computer system located at a site distant from the instrument can communicate wit the instrument, for example, through a wide area network or local area communication network or other suitable communication network. The system with the computer is programmed to automatically carry out steps of the methods herein and the requisite calculations. For embodiments that

A6

use mass-matched deoxyriboucleotides, a user enters the primer sequence or primer mass, the periodic reference mass and mass of an individual mass-matched deoxyonucleotide. These data can be directly entered by the user from a keyboard or from other computers or computer systems linked by network connection, or on removable storage medium such as a CD-ROM, minidisk (MD), DVD, floppy disk or other suitable storage medium. Next the user causes execution software that operates the system in which the mass spectrum of the extension products is generated. The Forced Mass Modulation software performs the steps of obtaining the masses of the fragments generated by the sequencing reaction and measured by the analytical instrument, and determining the identity of a nucleotide at any base position or the positional mass difference. The identity of the nucleotide at each base position is determined by comparing the calculated $M_{diff}[n]$ values to a database of previously calculated values of Mdiff for each of the chain terminating nucleotides.

Please replace the paragraph on pages 42, line 24, to page 43, line 3, with the following:

Hence, the periodicity is determined by the mass of the mass-matched nucleotide and the shift is the difference in location of a peak resulting from the chain terminator. For example, in Figure 2a, the lightest terminator is ddC, and the differential is O for C, 40 for G, 34 for A, 15 for T. The selected mass matched nucleotide has a mass of 310 Da. The primer in Figure 2a has a mass of 3327 Da and the first peak would be at 3600 if the first nucleotide in the extension product were C (0 shift). Since the first peak is at 3640, the shift is 40 Da. Therefore the first nucleotide is G, corresponding to a shift from the periodicity of 310 Da generated by the mass-matched nucleotides.

Please replace the paragraph on pages 45, line 10, to page 46, line 4, with the following:

EXAMPLE 1

Forced Mass Modulation using Mass-Matched Deoxynucleotides

For this implementation, a set of nucleotide analogs for the four bases in DNA are selected (Amersham Pharmacia Biotech) such that each base has exactly the same molecular weight, termed a *mass-matched* deoxynucleotide set. This is achieved by judiciously choosing chemical modifiers of the existing bases or by the using a universal base analog such as deoxyinosine, which can form base pairs with more than one of the natural bases. For this example, the mass of each A8 deoxynucleotide ("dN") in the mass-matched set has the arbitrarily selected value of 310 daltons, but any other value suffices. The sequencing reaction is performed as follows: 1) a primer is annealed to the target to be sequenced; 2) the resulting structure is subjected to an extension reaction using a suitable polymerase in the presence of the mass-matched nucleotide set and the four standard dideoxynucleotide terminators. The products and molecular masses of such a reaction are shown with a simulated mass spectrum in Figure 2a. The base periodicity is the mass of dN, or 310 daltons. The identity of a nucleotide at any base position is given by the *positional mass difference*, defined as the distance in daltons between the observed peak and the nearest *periodic reference mass*, which occurs every 310 daltons. In this example, the first periodic reference mass is defined as the (primer mass + ddC), or (3327 + 273) = 3600 daltons. The second periodic reference mass would be 3600 plus the base periodicity or (3600 + 310) = 3910, and so on. Expressed in terms of the base position n:

Please replace the paragraphs on page 49, lines 6-26, with the following:

Implementation of Forced Mass Modulation using pair-matched nucleotides is shown in Figure 4a-d. The basic requirement for this method is that the sequencing reaction products can be analyzed as double-stranded structures. A9 Briefly, the steps in the reaction are as follows: 1) A partially duplex hairpin primer with a 3′ overhang and a 5′ phosphate group is annealed and ligated to the single stranded target sequence. 2) The resulting partially duplex structure is subjected to a sequencing reaction using the pair-matched nucleotide set described above along with the set of mass-matched terminators (ddM). 3) The products from the sequencing reaction are exposed to a strict single strand-specific nuclease that

results in the production of blunt-ended hairpin structures ready for analysis by mass spectrometry. Figure 5a-b shows the products and molecular masses of the nuclease digestion along with a simulated mass spectrum.

A9

Because the reaction products are double-stranded, they are forced to assume a quasi-periodic distribution with a base periodicity of 617.4 daltons. The dotted regions on the spectrum shown in Figure 5b indicate the allowed mass ranges that can be occupied by the reaction products. The first periodic reference mass is at 10360 Da, which is the mass of the fully duplex hairpin primer plus a ddM:dC base pair. Expressing the periodic reference masses in terms of the base position n yields:

Please replace the table on page 50, lines 22-30, with the following:

A10

Terminator	*Nucleotide Analog*	*Mass*	*Base Pairing*	*Mass of Base Pair*
T	*5-Bromo-dideoxyuridine*	*353.1*	*5-Br-ddU:dA*	*666.3*
C	*5-Methyl-dideoxycytidine*	*287.2*	*5-Me-ddC: 7-deaza-dG*	*615.4*
A	*Dideoxyadenosine*	*297.2*	*ddA: dT*	*601.4*
G	*Dideoxyinosine*	*298.2*	*ddI: dC*	*587.4*

Please replace the paragraphs on page 52, line 8, to page 53, line 18, with the following:

A11

Forced Mass Modulation can be used to simplify the analysis of closely related sequence variants, as is required in the detection and scoring of single nucleotide polymorphisms. Figure 6a-d shows three sequence variants that differ from each other only at a single base position sequenced by a conventional Sanger reaction. The mass distribution of the reaction products is so complex that it can be uninterpretable, even if the base sequences of the variants are known *a priori*.

Figure 7a-d shows the same three variants sequenced by Forced Mass Modulation using mass-matched deoxynucleotides (dN = 310 Da) and the standard dideoxy terminators. The positions and identities of the single-nucleotide changes are immediately apparent from the mass spectrum. Since the masses of the sequencing reaction products are constrained to fall within the shaded regions of the spectrum in Figure 7b, it is possible to multiplex other sequences on the same spectrum.